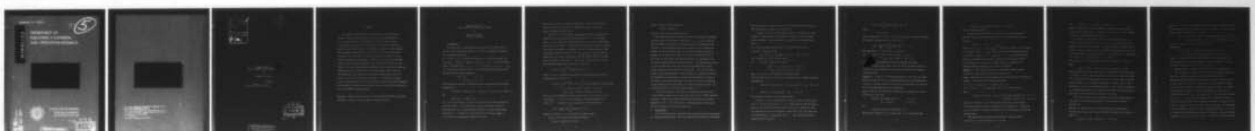
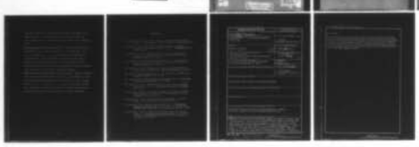
AD-A042 123 SYRACUSE UNIV N Y DEPT OF INDUSTRIAL ENGINEERING AND--ETC F/G 12/2
OPTIMAL CONTROL OF BIRTH AND DEATH PROCESSES AND QUEUES,(U)
1977 R F SERFOZO AF-AFOSR-2627-74

UNCLASSIFIED AFOSR-TR-77-0805 NL

1 OF 1
AD
A042123





END
DATE
FILMED
8-77

AFOSR-TR- 77-0805

AD A042123

5

DEPARTMENT OF INDUSTRIAL ENGINEERING AND OPERATIONS RESEARCH

COLLEGE OF ENGINEERING
SYRACUSE UNIVERSITY
SYRACUSE, NEW YORK 13210

(5)

Optimal Control of
Birth and Death Processes and

by

Richard F. Serfozo

Grant AFOSR-74-2627

Abstract

We consider a controlled birth and death process that moves as follows. Upon reaching a state i a pair of birth-death parameters (λ,μ) is selected from a prescribed set. Then the process remains in the state i until a birth or death occurs according to these parameters, at which time a pair of birth-death parameters is again selected. This is repeated indefinitely. A cost of $c(\lambda,\mu) + h(i)$ per unit time is incurred for selecting (λ,μ) when the process is in state i, and a reward is received for each birth. A policy is a rule for successively selecting the birth-death parameters as a function of the state of the process. We show, under some weak conditions, that there exist increasing optimal policies for both the discounted and average reward criteria. This means that it is optimal to increase the deaths and decrease the births as the state of the process increases. We show how to compute such an optimal policy for the case with two possible birth-death parameters. We then apply our results to the optimal control of the arrival and service rates in an M/M/1 queueing process.

Key Words: Markov decision processes, birth and death processes, queueing processes, stochastic control, monotone optimal policies.

Optimal Control of
Birth and Death Processes and Queues

by

Richard F. Serfozo
Syracuse University

1. Introduction

We shall study a controlled birth and death process that moves as follows. When the process arrives at a state i (a nonnegative integer) the following events occur.

(1) A pair (λ_a,μ_a) of birth-death parameters is selected from the set $\{(\lambda_1,\mu_1), \ldots, (\lambda_m,\mu_m)\}$. Think of (λ_a,μ_a) or $a \in \{1,\ldots,m\}$ as the action taken. We assume that $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_m > 0$, and $0 < \mu_1 \leq \mu_2 \leq \cdots \leq \mu_m$.

(2) The process remains in state i for a random time which has an exponential distribution with parameter

$$\lambda(i,a) = \begin{cases} \lambda_a & \text{if } i = 0 \\ \lambda_a + \mu_a & \text{if } i \geq 1. \end{cases}$$

Then the process jumps to a neighboring state according to the transition probabilities

$$q(i,a,i+1) = \lambda_a/(\lambda_a + \mu_a), \qquad q(i,a,i-1) = \mu_a/(\lambda_a + \mu_a) \quad \text{when } i \geq 1,$$

and

$$q(0,a,1) = 1 \quad \text{when } i = 0.$$

(3) A cost is incurred at a rate $c(a) + h(i)$ per unit time, during the sojourn in state i, for selecting the action a and being in state i. We assume that $c(a)$ is nondecreasing, and $h(i)$ is convex nondecreasing with $h(0) = 0$. A reward R is also received if the process jumps to $i + 1$: the R is a nonnegative reward for a birth.

This series of events is repeated indefinitely. Note that the birth-death parameters are selected at jump times of the process: they cannot be changed between jumps.

A policy f for successively choosing the birth-death parameters is defined to be a mapping from the state space $\{0,1,\ldots\}$ to the action space $\{1,\ldots,m\}$, with the interpretation that action f(i) is taken when the process is in state i. Each policy f, along with a rule for starting the process, determines a continuous time birth and death process whose birth-death parameters in state i are $(\lambda_{f(i)},\mu_{f(i)})$. We let Y_n and T_n denote the n-th state of the process, and the time at which the process jumps to state Y_n, respectively. The action in state Y_n is $a_n = f(Y_n)$. The discounted reward for the process is given by

$$W_f(i) = E_f\Big(\sum_{n=0}^{\infty} e^{-\beta T_n} g_\beta(Y_n,a_n) \mid X_o = i\Big),$$

where $\beta > 0$ is a continuous time discount factor and $g_\beta(i,a)$, the discounted gain in a sojourn, is

$$\begin{aligned} g_\beta(i,a) &= E_f\Big(e^{-\beta T_1} R\delta(Y_1,i+1) - \int_o^{T_1} (c(a) + h(i))e^{-\beta t}dt \mid Y_o=i,\ a_o=a\Big) \\ &= (R - c(a) - h(i))/(\beta + \lambda(0,a)) \qquad \text{if } i = 0 \\ &\quad (\lambda_a R - c(a) - h(i))/(\beta + \lambda(i,a)) \quad \text{if } i \geq 1. \end{aligned}$$

Here T_1 is the exponential sojourn time in state i, the $\delta(i,j) = 1$ or 0 according as $i = j$ or $i \neq j$, and a reward ρ received at time t has a value $\rho e^{-\beta t}$. Similarly, the average reward for the process is

$$\psi_f(i) = \underline{\lim}_{t\to\infty}\, t^{-1} E_f\Big(\sum_{n=0}^{N_t} g_o(X_n,a_n) \mid X_o = i\Big),$$

where $N_t = \sup\{n : T_n \leq t\}$ is the number of jumps in time t.

A policy f* is called β-discounted optimal if

$$W_{f^*}(i) = \sup_f W_f(i) \quad \text{for all } i,$$

and f* is called average optimal if

$$\psi_{f*}(i) = \sup_f \psi_f(i)$$

The aim is to find such optimal policies.

This controlled birth and death process is a continuous time Markov decision process with bounded sojourn rates $\lambda_a + \mu_a$. Such continuous time processes are equivalent to simpler discrete time Markov decision processes. This equivalence, which was originally used by Howard and Veinott (for processes with finite state spaces) and more recently by Lippman [5], is discussed in [9]. We use this equivalence herein to show that the controlled birth and death process is equivalent to the controlled random walk that we studied in [10]. Then applying the results in [10], we show that there exist increasing discounted and average optimal policies for the birth and death process. (We use increasing herein to mean nondecreasing.) This says that it is optimal to increase the probability of deaths and to decrease the probability of births as the state of the process increases. The results in [10] for computing average optimal policies also apply to the birth and death process. We illustrate this for a two action problem (when m = 2).

In the last section of this paper, we show how our results apply to the optimal control of the arrival and service rates of M/M/1 queueing processes. Some of the results herein have been derived by different approaches in [1] - [8], [11] and [12]. Bibliographies on the optimal control of queues are in [4], [11] and [12].

2. Main Results

We first show that the controlled birth and death process is equivalent to a controlled random walk. Next we establish the existence of increasing discount and average optimal policies for the birth and death process. Then we show how to compute one such policy.

We shall use the notation introduced above. In addition, we shall consider a random walk on the nonnegative integers (as in [10]) which moves as follows. Upon arriving at a location i the following events occur.

(i) A pair of probabilities (p_a, q_a) are selected from the set $\{(p_1,q_1),\ldots,(p_m,q_m)\}$, where

$$p_k = \lambda_k/\Lambda,\ q_k = \mu_k/\Lambda \text{ and } \Lambda = \lambda_1 + \mu_m \qquad \text{for } 1 \leq k \leq m.$$

(ii) A reward $r_\alpha(i,a)$ is received, where

$$r_\alpha(i,a) = \alpha g_\beta(i,a)(\beta + \lambda(i,a))\Lambda^{-1} \tag{1}$$

$$= \begin{cases} (R - c(a) - h(i))/(\beta + \Lambda) & \text{for } i = 0 \\ (\lambda_a R - c(a) - h(i))/(\beta + \Lambda) & \text{for } i \geq 1, \end{cases}$$

and $\alpha = \Lambda/(\beta + \Lambda)$ is a discrete time discount factor.

(iii) The next state of the walk is determined by the following transition probabilities

$$p(i,a,i+1) = p_a,\ p(i,a,i) = 1-p_a-q_a,\ p(i,a,i-1) = q_a \qquad \text{for } i \geq 1,$$

and

$$p(0,a,1) = p_a \text{ and } p(0,a,0) = 1-p_a \qquad \text{for } i = 0.$$

This series of events is repeated indefinitely.

A policy f for this random walk is a function from the state space $\{0,1,\ldots\}$ to the action space $\{1,\ldots,m\}$. (These policies are the same as those for the birth and death process.) A policy f, along with a rule for starting the process, determines a random walk $\{X_n : n \geq 0\}$, where the n-th action taken is (p_a, q_a) when $f(X_n) = a$. The expected discounted reward for this process is

$$V_f(i) = E_f\Big(\sum_{n=0}^{\infty} \alpha^n r_\alpha(X_n, f(X_n)) \mid X_o = i\Big),$$

where

$$\alpha = \Lambda/(\beta + \Lambda)$$

The V_f exists and $-\infty \leq V_f(i) < \infty$, since the $r(i,a)$ is bounded from above. The expected average return for the process is

$$\phi_f(i) = \varliminf_{n\to\infty} n^{-1} E_f\Big(\sum_{k=0}^{n-1} r_1(X_k, f(X_k)) \mid X_o = i\Big).$$

Note that when $\alpha = 1$ in (1) the $\beta = 0$, and so

$$r_1(i,a) = \begin{cases} (R - c(a) - h(i))\Lambda^{-1} & \text{for } i = 0 \\ (\lambda_a R - c(a) - h(i))\Lambda^{-1} & \text{for } i \geq 1. \end{cases} \tag{2}$$

Discount and average optimal policies are defined as before.

The following result asserts that the birth and death process is equivalent to this random walk in the sense that they have identical optimal policies.

Theorem 2.1. A policy is β-discounted optimal for the birth and death process if and only if it is α-discounted optimal for the random walk. A policy is average optimal for the birth and death process if and only if it is average optimal for the random walk.

Proof. Note that the transition probabilities and rewards of the two processes are such that

$$p(i,a,j) = \begin{cases} \lambda(i,a)q(i,a,j)\Lambda^{-1} & \text{if } i \neq j \\ 1 - \lambda(i,a)\Lambda^{-1} & \text{if } i = j, \end{cases}$$

and

$$r_\alpha(i,a) = g_\beta(i,a)(\beta + \lambda(i,a))/(\beta + \Lambda).$$

Then from [9, Theorem 1.1] it follows that if f is any policy then

$$W_f(i) = V_f(i) \text{ and } \psi_f(i) = \Lambda\phi_f(i) \qquad \text{for all } i.$$

Thus the assertions follow.

Our next result establishes the existence of increasing optimal policies for our birth and death process. An increasing policy is of the form

$$f(i) = a \qquad \text{if } i_a \leq i < i_{a+1},$$

where $0 = i_1 \leq i_2 \leq \cdots \leq i_m \leq i_{m+1} = \infty$. Under this policy, if the process is in state i and $i_a \leq i < i_{a+1}$, then the birth-death parameters (λ_a, μ_a) are selected. Since $\lambda_1 \geq \cdots \geq \lambda_m$ and $\mu_1 \leq \cdots \leq \mu_m$, then the selected death rate increases as the state i increases, and the selected birth rate decreases as i increases. In other words, the probability of backward movement increases as the state i increases.

Theorem 2.2. *There exist increasing discounted and average optimal policies for controlling the birth and death process.*

Proof. Consider the controlled random walk we defined above for discounted rewards. It satisfies the following conditions:

(3) $p_1 \geq \cdots \geq p_m$, $q_1 \leq \cdots \leq q_m$ and $p_1 + q_m \leq 1$.

(4) $r_\alpha'(i,1) \leq \cdots \leq r_\alpha'(i,m)$ and $r_\alpha'(i,1) \geq r_\alpha'(i+1,m)$ for all i,

where

$$r_\alpha'(i,a) = r_\alpha(i+1,a) - r_\alpha(i,a) = \begin{cases} -[(1-\lambda_a)R + h(1)]/(\beta + \Lambda) & \text{if } i = 0 \\ -(h(i+1) - h(i))/(\beta + \Lambda) & \text{if } i \geq 1. \end{cases}$$

Then by [10, Theorem 2.1] it follows that there exists an increasing α-discounted optimal policy for controlling the random walk. This policy is also β-optimal, according to Theorem 2.2, for controlling the birth and death process.

Now consider the controlled random walk for average rewards (recall (2)). It also satisfies (3) - (4) and

(5) $\qquad r_1(i,1) \geq \cdots \geq r_1(i,m) \qquad$ for all i.

Then by [10, Theorem 6.1], there exists an increasing average optimal policy for the random walk. By Theorem 2.2, this policy is also average optimal for the birth and death process.

In Theorem 2.2, we assumed that the birth-death parameters are selected at each jump from a set $\{(\lambda_1,\mu_1),\ldots,(\lambda_m,\mu_m)\}$ which is independent of the state of the process. Suppose instead, that when the process is in state i, then a pair of birth-death parameters $(\bar\lambda(i,a),\bar\mu(i,a))$ is selected from a set $\{(\bar\lambda(i,1),\bar\mu(i,1)),\ldots,(\bar\lambda(i,m),\bar\mu(i,m))\}$. Assume that

(6) $\bar\lambda(i,1) \geq \cdots \geq \bar\lambda(i,m)$, $\bar\mu(i,1) \leq \cdots \leq \bar\mu(i,m)$,

(7) $d'(i,1) \leq \cdots \leq d'(i,m) \leq 0$, and $d'(i,1) \geq d'(i+1,m)$ for all i,

where

$$d(i,a) = \mu(i,a) - \lambda(i,a) \text{ and } d'(i,a) = d(i+1,a) - d(i,a)$$

This controlled birth and death process is equivalent to the random walk in Section 3 of [10]. The proof of this is the same as that for Theorem 2.1. From this equivalence, along with [10, Theorem 3.1] and an average reward analog of it, it follows that Theorem 2.2 holds. This is also discussed in [1] just for discounted rewards. Other analogs of Theorem 2.2 for decreasing policies, or for finite time horizons, can be obtained in the same way from Theorem 2.3 or Theorem 9.1 in [10].

The results in [10] on the computation of average optimal policies also apply to birth and death processes. We illustrate this for a special case.

Theorem 2.4. Suppose the birth and death process has two possible parameter pairs (λ_1,μ_1) and (λ_2,μ_2) and that the reward for a sojourn in state i is

$$g_o(i,a) = (-c(a) - ih)/\lambda(i,a) \qquad \text{for all } a,$$

where $h > 0$. Then it is average optimal to select (λ_1, μ_1) when the state of the process is below n^* and to use (λ_2, μ_2) otherwise. Here n^* is the smallest integer for which $D_n \geq 0$, where

$$D_n = \begin{cases} n + b\rho_1^{\,n} + b - \rho_2(c(2) - c(1))/(h(\rho_1 - \rho_2)(1-\rho_2)) & \text{if } \rho_1 \neq 1 \\ n^2 + n(1+\rho_2)/(1-\rho_2) - 2\rho_2(c(2) - c(1))/(h(\rho_1-\rho_2)) & \text{if } \rho_1 = 1 \end{cases}$$

and $\rho_a = \lambda_a/\mu_a$ and $b = (\rho_1 - \rho_2)/((1-\rho_1)(1-\rho_2))$.

Furthermore

$$n^* \leq \begin{cases} \rho_2(c(2) - c(1))/(h(\rho_1-\rho_2)(1-\rho_2)) & \text{if } \rho_1 \neq 1 \\ [2\rho_2(c(2) - c(1))/(h(\rho_1-\rho_2))]^{1/2} & \text{if } \rho_1 = 1. \end{cases}$$

Proof. The policy described in this theorem is average optimal, by [10, Corollary 7.2], for the random walk, Thus, by Theorem 2.2, it is average optimal for the birth and death process.

3. Optimal Control of Arrival and Service Rates in an M/M/1 Queue

The following are examples of controlled birth and death processes which have monotone optimal policies as we discussed above.

M/M/1 Queue with a Controlled Service Rate. Suppose an M/M/1 queue has a fixed arrival rate λ and its service rate is controlled as follows. At each service completion or customer arrival, the number of customers in the system is observed. Based on this number a service rate μ_a is selected from the set $\{\mu_1, \ldots, \mu_m\}$, where the μ's are subscripted so that $0 < \mu_1 < \ldots < \mu_m$. A cost $c(a)$ per unit time is charged for using μ_a, and a cost $h(i)$ per unit time is charged for holding i customers in the system. A reward R is also received from each customer. We assume that $c(a)$ is increasing, and $h(i)$ is convex increasing and $h(0) = 0$. This is a controlled birth and death process as in Theorem 2.2 with birth-death parameters $\{(\lambda, \mu_1), \ldots, (\lambda, \mu_m)\}$. Thus it is optimal (for both discounted and average rewards) to increase the service rate as the number of customers increases. This was proved in [3], and similar results for finite length queues are discussed in [2], [8] and some of the references in [4].

M/M/1 Queue with a Controlled Arrival Rate. Suppose an M/M/1 queue has a fixed service rate μ, and the arrival rate λ_a is selected from a set $\{\lambda_1,\ldots,\lambda_m\}$, where $\lambda_1 > \ldots > \lambda_m > 0$, at each service completion or customer arrival. Costs $c(a)$ and $h(i)$ are incurred as above, and a reward R is received from each customer. Then by Theorem 2.2 it is optimal to decrease the arrival rate as the number of customers increases. This was proved in [7]: see [6] for finite queue lengths.

M/M/1 Queue with Controlled Arrival and Service Rates. Suppose in an M/M/1 queue that the arrival and service rate pair (λ_a,μ_a) is selected, at each service completion and customer arrival, from a set $\{(\lambda_1,\mu_1),\ldots,(\lambda_m,\mu_m)\}$ where $\lambda_1 \geq \ldots \geq \lambda_m > 0$ and $0 < \mu_1 \leq \ldots \leq \mu_m$. With the costs $c(a)$ and $h(i)$, and reward R, as above, it is optimal to increase the service rate and decrease the arrival rate as the number of customers increases.

References

[1] Albright, S.C. and W. Winston (1977) A birth death model of advertising. Technical report. School of Business, Indiana University.

[2] Beja, A. and A. Teller (1975) Relevant policies for Markovian queueing systems with many types of service. Management Science, 21, 1049-1054.

[3] Crabill, T. (1972) Optimal control of a service facility with variable exponential service time and constant arrival rate. Management Science, 18, 560-566.

[4] Crabill, T., Gross, D. and M. Magazine (1977) A classified bibliography of research on optimal design and control of queues. Operations Research, 25, 219-232.

[5] Lippman, S. (1975) Applying a new device in the optimization of exponential queueing systems. Operations Research, 23, 687-710.

[6] Low, D. (1974) Optimal dynamic pricing policies for an M/M/s queue. Operations Research, 22, 545-561.

[7] Low, D. (1974) Optimal pricing for an unbounded queue. IBM Journal of Research and Development, 18, 290-302.

[8] Sabeti, H. (1973) Optimal selection of service rates in queueing with different costs. J. Operations Research Soc. of Japan, 16, 15-35.

[9] Serfozo, R. (1976) An equivalence between continuous and discrete time Markov decision processes. Technical Report, IE & OR Department, Syracuse University.

[10] Serfozo, R. (1977) Optimal control of random walks. Technical Report, IE & OR Department, Syracuse University.

[11] Sobel, M. (1974) Optimal operation of queues. In Mathematical Methods in Queueing Theory. Lecture notes in Economics and Mathematical Systems, Vol. 98 (Springer, New York) pp. 231-261.

[12] Stidham, S. and N.U. Prabhu (1974) Optimal control of queueing systems. In Mathematical Methods in Queueing, Lecture notes in Economics and Mathematical Systems, Vol. 98. (Springer, New York), pp. 263-294.

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: 19 18 AFOSR-TR-77-0805
2. GOVT ACCESSION NO.:
3. RECIPIENT'S CATALOG NUMBER:
4. TITLE (and Subtitle): 6 OPTIMAL CONTROL OF BIRTH AND DEATH PROCESSES AND QUEUES,
5. TYPE OF REPORT & PERIOD COVERED: Interim
6. PERFORMING ORG. REPORT NUMBER:
7. AUTHOR(s): 10 Richard F./Serfozo
8. CONTRACT OR GRANT NUMBER(s): 15 AF-AFOSR-~~74~~-2627-74
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Syracuse University, Dept of Industrial Engr & Operations Research, Syracuse, NY 13210
10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 61102F 16 2304/A5 17 A5
11. CONTROLLING OFFICE NAME AND ADDRESS: Air Force Office of Scientific Researc/NM, Bolling AFB, Washington, DC 20332
12. REPORT DATE: 11 1977
13. NUMBER OF PAGES: 9 12 16p.
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office):
15. SECURITY CLASS. (of this report): UNCLASSIFIED
15a. DECLASSIFICATION/DOWNGRADING SCHEDULE:

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Markov decision processes, birth and death processes, queueing processes, stochastic control, monotone optimal policies

lambda, mu

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

We consider a controlled birth and death process that moves as follows. Upon reaching a state i a pair of birth-death parameters (λ,μ), is selected from a prescribed set. Then the process remains in the state i until a birth or death occurs according to these parameters, at which time a pair of birth-death parameters is again selected. This is repeated indefinitely. A cost of $c(\lambda,\mu) + h(i)$ per unit time is incurred for selecting $(\lambda\mu)$ when the process is in state i, and a reward is received for each birth. A policy is a rule for successively selecting the birth-death parameters as a function of

next page

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFEID

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

lambda, mu 409184.

20. Abstract

the state of the process. We show, under some weak conditions, that there exist increasing optimal policies for both the discounted and average reward criteria. This means that it is optimal to increase the deaths and decrease the births as the state of the process increases. We show how to compute such an optimal policy for the case with two possible birth-death parameters. We then apply our results to the optimal control of the arrival and service rates in an M/M/1 queueing process.